Channel Inversion and Regularization Revisited

Signal Processing 121 (2016) 70–80

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Channel inversion and regularization revisited ☆

S. Morteza Razavi *, Tharmalingam Ratnarajah

Institute for Digital Communications, The University of Edinburgh, UK

ARTICLE INFO

Article history:
Received 15 February 2015
Received in revised form
6 October 2015
Accepted 15 October 2015
Available online 2 November 2015

Keywords:
Channel inversion
Imperfect CSI
Multiantenna downlink
Regularized channel inversion

ABSTRACT

In multiuser multiple-input-single-output (MISO) downlink, linear precoders like channel inversion (CI) and regularized CI (RCI) are more desirable than their nonlinear counterparts due to their reduced complexity. To achieve the full benefits of linear precoding, the availability of perfect channel state information (CSI) at base stations (BSs) is necessary. Since in practice, having access to perfect CSI is not pragmatic, in this paper, we evaluate the performance of CI and RCI under a generalized, imperfect CSI model where the variance of the channel estimation error depends on the signal-to-noise ratio (SNR) and thus covers digital and analog feedbacks as two special cases. Then, based on this imperfect CSI model, we quantify the asymptotic mean loss in sum rate and the achievable degrees of freedom (DoFs) by deriving the received signal-to-interference-plus-noise ratio (SINR) of each user. For example, it is shown that the achievable DoF is directly related to the SNR exponent of the channel estimation error variance. Also, two asymptotic gaps to capacity for the analog feedback are derived: mean loss in sum rate and power loss. In addition, we propose an adaptive RCI technique by deriving an appropriate regularization parameter as a function of the error variance and without imposing any restrictions on the number of users or antennas. It is shown that in the presence of CSI mismatch, while the comparative improvement of the standard RCI to CI becomes negligible, the adaptive RCI compensates this degraded performance of the standard RCI without introducing any extra computational complexity.

1. Introduction

1.1. Background

Thanks to multiple antennas, each base station (BS) is able to communicate to more than one user simultaneously at the expense of intra-cell interference. This scenario is referred to as multiuser multiple-input-single-output (MISO) downlink. One effective way to presuppress this interference is to deploy linear precoding at the BS. Linear precoders are well-acknowledged techniques owing to their reduced complexity compared to the nonlinear precoding techniques such as dirty paper coding (DPC) [1], vector perturbation [2,3], and Tomlinson–Harashima [4]. The least complex and the most prevalent technique is channel inversion (CI) [5], which is a linear precoding technique that yields reasonable performance. For instance, it has been shown that CI precoding, while generally suboptimal, can achieve the same asymptotic sum capacity as DPC does, when the number of users unlimitedly increases [6]. Nevertheless, in a case where the number of antennas at the BS is equal to the total number of single-antenna users and both are finite, it has been shown that with proportionally increasing the number of transmit and receive antennas, the bit error rate (BER) of each user, caused by deploying CI precoding, deteriorates. Also in this

☆ This work was supported in part by the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission under FET-Open grant number: HíATUS-265578.

* Corresponding author.
E-mail addresses: Morteza.Razavi@ed-alumni.net (S.M. Razavi), T.Ratnarajah@ed.ac.uk (T. Ratnarajah).

http://dx.doi.org/10.1016/j.sigpro.2015.10.016
0165-1684/

case, while the sum capacity linearly increases, the achievable sum rate of CI fails to do so.

Regularized channel inversion (RCI) [5], on the other hand, improves the performance of CI such that with increasing the number of antennas, the BER of each user remains fixed at low-to-intermediate signal-to-noise ratios (SNRs) and slightly improves at high SNRs. Plus, by using RCI, the sum rate has now a linear growth with the number of transmit antennas. Moreover, even by considering a fixed number of antennas at the BS, RCI achieves higher throughput than CI at low-to-intermediate SNRs.

Nevertheless, the asymptotic optimality of CI precoding [6] and the superiority of the standard RCI to CI [5] are subject to the availability of perfect channel state information (CSI) at the BS, which is a very stringent requirement in practice. Hence, performance analysis of CI and RCI under a generalized CSI mismatch model is of particular interest and is thus going to be addressed in this paper.

Many efforts have been accomplished to outline the performance of linear precoding in the presence of channel imperfections. Most of these works are related to the quantized feedback strategies, e.g., [7–16], and some of them considered the performance analysis in reciprocal channels, e.g., [17–21]. Although these works mainly considered a very specific scenario for imperfect CSI, i.e., CSI feedback or reciprocal channels, there are also few number of literature that analyzed the performance of linear precoding under a rather generalized CSI mismatch model by deriving some asymptotic bounds [22–25]. For example, in [7], it has been shown that in MIMO broadcast channels with finite-rate feedback, full degrees of freedom (DoFs) can be achieved if the number of fedback bits scales fast enough with signal-to-noise ratio (SNR). Also, [23] derived upper and lower bounds on the achievable rates of zero-forcing (ZF) beamforming under pilot-based channel estimation, i.e., analog feedback, and explicit channel state feedback, i.e., digital feedback.

1.2. Contributions

Compared to the most of the previous works on linear precoding under imperfect CSI, the proposed CSI mismatch model in this work is generalized to the case where the variance of the channel estimation error is considered to be a function of the SNR. This provides a very tractable and versatile CSI mismatch model that can cover a variety of distinct scenarios like perfect CSI, analog and digital feedbacks. Then under this imperfect CSI model, we derive novel bounds regarding the asymptotic mean loss in sum rate and the achievable DoFs.

Moreover, we consider the performance improvement of RCI precoding by deriving an optimum regularization parameter. So far, this problem has been treated in two distinguishable ways: In the first category of literature like [8,9], the regularization parameter has been derived only for the case of digital feedback where it has been further assumed that the perfect CSI and the channel estimation error are not independent of each other. In the second category of literature like [5,24–27], the regularization parameter has been derived under the assumption of asymptotically large number of transmit and receive antennas. However, the proposed regularization parameter in this work is derived in a distinct manner from the previously mentioned works such that it is amenable to our generalized CSI mismatch model. Furthermore, no restrictions on the number of transmit and receive antennas are imposed.

The proposed scheme is dubbed adaptive RCI since the sought regularization parameter is based on the knowledge of the channel estimation error variance which is possible to be known in advance, owing to the channel dynamics and channel estimation schemes. Then, we compare the performance of adaptive RCI with standard RCI proposed in [5]. First, it is shown that under perfect CSI, adaptive RCI boils down to the standard RCI, which implies on the generality of the derived regularization parameter which covers the cases of perfect and imperfect CSI. More importantly, it is shown that in the presence of CSI mismatch, while the performance improvement of standard RCI compared to CI becomes negligible, the proposed adaptive RCI compensates this degraded performance of the standard RCI by achieving higher sum rates and lower BERs. In particular, it is demonstrated that under digital feedback, while the sum rate and the BER of standard RCI experience a nonmonotonic trend, those of the proposed adaptive RCI manifest a monotonic behavior. This further implies that under digital feedback and at high SNRs, the performance of standard RCI is the same as that of CI whereas adaptive RCI distinguishably outperforms both of them.

1.3. Paper organization

We present the system model under perfect and imperfect CSI in Section 2. In Section 3, we quantify novel bounds on the asymptotic performance of linear downlink precoding. Section 4 deals with the performance analysis of standard RCI under imperfect CSI by deriving the output SINR of each user. In Section 5, an adaptive RCI technique is proposed by deriving an optimum regularization parameter. In Section 6, we use numerical simulations to corroborate the undergone analyses in this paper, and finally Section 7 contains conclusions.

1.4. Notations

Throughout the paper, a is a scalar, $\mathbf{a}$ is a vector, and $\mathbf{A}$ is a matrix. The superscript $(\cdot)^{\mathsf{H}}$ represents the Hermitian transpose. $\mathbb{E}\{\cdot\}$ and $\mathrm{Tr}[\cdot]$ are the expectation and trace operators, respectively. While $\|\cdot\|_2$ denotes the vector 2-norm, $[\cdot]_{\ell,\ell}$ designates the ℓth diagonal element of a square matrix.

2. Preliminaries

2.1. System model

We consider a multiuser downlink scenario where an N-antenna transmitter communicates with mobile terminals (MTs) with M receive antennas in total, as depicted in Fig. 1. Since no signal processing treatment is going to be considered at each MT, the system configuration is irrespective to whether the receive antennas cooperate or not,

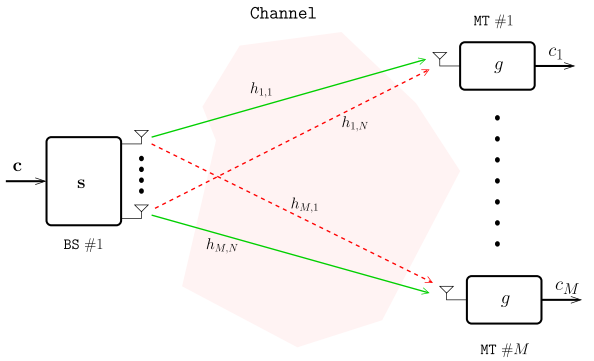

Fig. 1. Single-cell broadcast channel where dash red arrows represent intra-cell interference while solid green arrows denote desired links. h_{kj} is the time-variant channel response between the jth transmit antenna of the BS and the kth MT.

therefore the total number of receive antennas can belong to one user or be shared by several users; however, as purely transmitter-based precoders are most useful with single antenna receivers, we consider single-antenna MTs in the following, which is also consistent with most of the references in this work. Without loss of generality, we assume that all single-antenna users are homogeneous and experience independent fading. The received signals of all users can be collectively expressed by

$$\mathbf{y} = \sqrt{P}\mathbf{H}\mathbf{s} + \mathbf{z} \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{M\times 1}$, P is the transmit power, and $\mathbf{H} \in \mathbb{C}^{M\times N}$ denotes the channel from N-antenna transmitter to M single-antenna users such that the magnitude of channel coefficients is bounded away from zero and infinity. We further assume that elements of $\mathbf{H}$ can be modeled by independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and unit variance, i.e., $\mathrm{vec}\,(\mathbf{H}) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I})$, $\mathbf{s} \in \mathbb{C}^{N\times 1}$ is the transmitted signal from the BS, and $\mathbf{z} \in \mathbb{C}^{M\times 1}$ is the circularly symmetric additive white Gaussian noise with zero mean and variance σ^2, i.e., $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2\mathbf{I})$. We further assume that the transmitted signal $\mathbf{s}$ in (1) can be expressed as $\mathbf{s} = g\mathbf{\Psi}\mathbf{c}$. Similar to [5,27–29], we consider g as the scaling factor that ensures transmit power constraint i.e., $\mathbb{E}\left\{\|\mathbf{s}\|_2^2\right\} = 1$. $\mathbf{\Psi}$ is the precoding matrix and $\mathbf{c}$ represents the vector containing the symbols chosen from a desired constellation and since we assume i.i.d. Gaussian input signaling, we have $\mathbb{E}\{\mathbf{c}\mathbf{c}^{\mathsf{H}}\} = \mathbf{I}$. We also define the nominal SNR as $\rho = P/\sigma^2$. Note that although the concept of regularization is most beneficial for the case of equal number of transmit and receive antennas [5,24,27,28], without loss of generality, we assume $M \le N$.

2.2. Imperfect CSI model

Unlike some of the earlier works, where the perfect CSI, viz. $\mathbf{H}$, is typically considered to be dependent on the channel estimation error, here, we model the imperfect CSI as [29,30,31]

$$\widehat{\mathbf{H}} = \mathbf{H} + \mathbf{E} \tag{2}$$

where the actual channel matrix $\mathbf{H}$ is thought to be independent of channel measurement error $\mathbf{E}$. We further consider $\mathbf{E}$ as a Gaussian matrix consisting of i.i.d. elements with mean zero and variance τ, i.e., [30]

$$\mathrm{vec}(\mathbf{E}) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tau\mathbf{I}) \quad \text{with } \tau \triangleq \beta\rho^{-\alpha}, \quad \beta > 0,\ \alpha \ge 0 \tag{3}$$

In this case, the error variance can depend on the SNR ($\alpha \neq 0$) or be independent of that ($\alpha = 0$). In particular, perfect CSI is regained by setting $\tau = 0$. Notice the variance model in (3) is versatile since it is potentially able to accommodate a variety of distinct scenarios. More specifically, τ can be interpreted as a parameter that captures the quality of the channel estimation which is possible to be known a priori, depending on the channel dynamics and channel estimation schemes, see e.g., [32] and references therein. The four cases of this error variance model can be described as follows:

- *CSI feedback*: In this case, the channel matrix can be estimated by pilot transmissions in the downlink. Then, a quantized version of this channel estimate is sent back to the BS through a dedicated feedback link. This way, the imperfect CSI will be mostly dominated by the errors caused through quantization and feedback delay, which can eventually result in outdated CSI at the BS if the channel coherence time is smaller than the feedback delay. Since channel coherence time and the resolution of quantizer do not depend on ρ, the channel estimation error variance τ becomes independent of ρ as well. This case is captured by setting $\alpha = 0$.
- *Reciprocal channels*: This case represents the reciprocal systems like time division duplex where uplink and downlink channels are identical. The downlink channel can thus be estimated through pilots sent over the uplink channel and the channel measurement error $\mathbf{E}$ depends on the noise level at the BS as well as the pilot power. If the pilot power proportionally increases with P, the channel estimation error scales inversely with increasing ρ. This case is modeled by setting $\alpha = 1$.
- $0 < \alpha < 1$: This may be the case where the BS and mobile transmit powers are not in the same range, or of the same order, such that the feedback power is much smaller than the feedforward power.
- $\alpha > 1$: This may be the case where the BS and mobile transmit powers are in the same range but the feedback power is attenuated in comparison with the feedforward power.

To facilitate the performance analysis of CI and RCI under CSI mismatch model in (2), it is more appropriate to have the statistical properties of $\mathbf{H}$ conditioned on $\widehat{\mathbf{H}}$. In this case, conditioned on $\widehat{\mathbf{H}}$, $\mathbf{H}$ has a Gaussian distribution with mean $\widehat{\mathbf{H}}/(1+\tau)$ and statistically independent elements of variance $\tau/(1+\tau)$ [33], i.e.,

$$\mathbf{H} = \frac{1}{1+\tau}\widehat{\mathbf{H}} + \check{\mathbf{H}} \tag{4}$$

where the auxiliary random matrix

$$\mathrm{vec}\left(\check{\mathbf{H}}\right) \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \frac{\tau}{1+\tau}\mathbf{I}\right)$$

is statistically independent of $\widehat{\mathbf{H}}$.

3. Asymptotic performance of channel inversion

In this section, we first derive the output SINR of each user when CI is deployed at the BS. We then derive novel bounds on the asymptotic mean loss in sum rate and the achievable DoFs when the BS is in possession of imperfect CSI. Consequently, we assume that the channel estimate $\widehat{\mathbf{H}}$ is only available, and the signal preprocessing at the BS is thus going to be done upon the knowledge of $\widehat{\mathbf{H}}$.

3.1. Channel inversion under perfect CSI

When the perfect channel state information is available at the BS, the transmitted signal can be represented as

$$\mathbf{s}_{\mathrm{CI}} = g_{\mathrm{CI}}\boldsymbol{\Psi}_{\mathrm{CI}}\mathbf{c} \tag{5}$$

where the precoding matrix is $\boldsymbol{\Psi}_{\mathrm{CI}} = \mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}$ and the scaling factor can be defined as [5]

$$g_{\mathrm{CI}} = \frac{1}{\sqrt{\mathrm{Tr}\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right]}} \tag{6}$$

In this case, the received signal can be shown as

$$\mathbf{y}_{\mathrm{CI}} = \sqrt{P}g_{\mathrm{CI}}\mathbf{H}\mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\mathbf{c}+\mathbf{z} = \sqrt{P}g_{\mathrm{CI}}\mathbf{c}+\mathbf{z} \tag{7}$$

and consequently the unified output SINR of each user is equal to

$$\eta_{\mathrm{CI}} = \frac{Pg_{\mathrm{CI}}^2}{\sigma^2} \tag{8}$$

With respect to the fact that

$$\mathrm{Tr}\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right] = \sum_{\ell=1}^{M}\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right]_{\ell,\ell}$$

the output SINR of the ℓth user can be shown as [6,28]

$$\dot{\eta}_{\mathrm{CI}} = \frac{P}{M\sigma^2\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right]_{\ell,\ell}} \tag{9}$$

Without loss of generality and to avoid cumbersome formulation, and also to simplify the analysis within this subsection and also the next one, we use the unified output SINR instead of the output SINR of the ℓth user, since this interchangeability does not compromise the validity of the asymptotic performance analysis at high SNRs. The achievable sum rate under perfect CSI by considering the unified output SINR can thus be expressed as [5]

$$R_{\mathrm{Perfect\ CSI}} = M\log_2(1+\eta_{\mathrm{CI}}) = M\log_2\left(1+\frac{Pg_{\mathrm{CI}}^2}{\sigma^2}\right) \tag{10}$$

and the total achievable DoF is equal to [7]

$$\begin{aligned} D_{\mathrm{Perfect\ CSI}} &= \lim_{P\to\infty}\frac{\mathbb{E}_{\mathbf{H}}\{R_{\mathrm{Perfect\ CSI}}\}}{\log_2 P} \\ &= \lim_{P\to\infty}\frac{\mathbb{E}_{\mathbf{H}}\left\{M\log_2\left(1+\frac{Pg_{\mathrm{CI}}^2}{\sigma^2}\right)\right\}}{\log_2 P} = M \end{aligned} \tag{11}$$

3.2. Channel inversion under imperfect CSI

In the presence of the imperfect CSI at the BS, the precoding matrix can now be defined as $\widehat{\boldsymbol{\Psi}}_{\mathrm{CI}} = \widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}$. Consequently in this case, the transmitted signal in (1) can be shown as

$$\widehat{\mathbf{s}}_{\mathrm{CI}} = \widehat{g}_{\mathrm{CI}}\widehat{\boldsymbol{\Psi}}_{\mathrm{CI}}\mathbf{c} \tag{12}$$

where the scaling factor is equal to

$$\widehat{g}_{\mathrm{CI}} = \frac{1}{\sqrt{\mathrm{Tr}\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]}} \tag{13}$$

Therefore the received signal in (1) can be represented by

$$\begin{aligned} \widehat{\mathbf{y}}_{\mathrm{CI}} &= \sqrt{P}\widehat{g}_{\mathrm{CI}}\mathbf{H}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\mathbf{c}+\mathbf{z} \\ &\overset{①}{=} \sqrt{P}\widehat{g}_{\mathrm{CI}}\left(\frac{1}{1+\tau}\widehat{\mathbf{H}}+\check{\mathbf{H}}\right)\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\mathbf{c}+\mathbf{z} \\ &= \underbrace{\frac{\sqrt{P}\widehat{g}_{\mathrm{CI}}}{1+\tau}\mathbf{c}}_{\text{desired term}} + \underbrace{\sqrt{P}\widehat{g}_{\mathrm{CI}}\check{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\mathbf{c}+\mathbf{z}}_{\text{interference plus noise term}} \end{aligned} \tag{14}$$

where ① follows from (4). Note that in this case and as revealed in (14), to have an unbiased detection, the received signals should be scaled back by $(1+\tau)/\widehat{g}_{\mathrm{CI}}$. To further proceed, we consider the following lemma:

Lemma 1. *If $\mathbf{A} \in \mathbb{C}^{M\times N}$ represents a Gaussian matrix with i.i.d. elements of mean zero and variance a, then $\mathbb{E}\left\{\mathbf{A}\mathbf{A}^{\mathrm{H}}\right\} = aN\,\mathbf{I}$.*

Proof. Since $\mathbf{A}$ is a Gaussian matrix, we have $\mathrm{vec}(\mathbf{A}) \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, a\mathbf{I})$. In other words, if $\mathbf{a}^{\mathrm{H}}$ represents an arbitrary column of $\mathbf{A}$, then $\mathbb{E}\{\mathbf{a}^{\mathrm{H}}\mathbf{a}\} = a\,\mathbf{I}$ [34]. However, since $\mathbf{A}$ has N independent columns, the claim follows.□

Note that throughout the paper, we assume that the noise and data vectors are independent of each other and are also independent of the actual channel matrix $\mathbf{H}$, which is consistent with [5]. Since $\mathbf{H}$ depends on both $\widehat{\mathbf{H}}$ and $\check{\mathbf{H}}$, the data and noise are likewise considered to be independent of $\widehat{\mathbf{H}}$ and $\check{\mathbf{H}}$. However, to make the output SINR of each user dependent only on the channel estimate $\widehat{\mathbf{H}}$, we additionally take the expectation over $\check{\mathbf{H}}$. This is also consistent with [8,35] wherein the expectation was taken over the auxiliary channel measurement error.

Therefore, by considering Lemma 1, we have

$$\mathbb{E}_{\check{\mathbf{H}}}\left\{\check{\mathbf{H}}^{\mathrm{H}}\check{\mathbf{H}}\right\} = \frac{M\tau}{1+\tau}\mathbf{I}$$

Also with respect to (13), it is straightforward to show that

$$\mathbb{E}_{\check{\mathbf{H}},\mathbf{c}}\left\{\left\|\sqrt{P}\widehat{g}_{\mathrm{CI}}\check{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\mathbf{c}\right\|_2^2\right\}=\frac{PM\tau}{1+\tau}$$

and consequently for a given realization of $\widehat{\mathbf{H}}$, and with respect to the normalizing factor, the unified output SINR of each user, as a function of τ, can be given by

$$\widehat{\eta}_{\mathrm{CI}}=\frac{P\widehat{g}_{\mathrm{CI}}^2}{P\tau(1+\tau)+\sigma^2(1+\tau)^2} \tag{15}$$

Since $\mathrm{Tr}\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]=\sum_{\ell=1}^{M}\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]_{\ell,\ell}$ and by considering (8)–(9), the output SINR of the ℓth user can be shown as

$$\ddot{\eta}_{\mathrm{CI}}=\frac{P}{M\left(P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]_{\ell,\ell}} \tag{16}$$

Note that by setting $\tau=0$, $\ddot{\eta}_{\mathrm{CI}}$ boils down to $\dot{\eta}_{\mathrm{CI}}$ in (9) which is the instantaneous output SINR of the ℓth user under perfect CSI.

Consequently and by considering the unified output SINR in (15), the achievable sum rate of CI under the imperfect CSI can be represented by

$$\begin{aligned}R_{\mathtt{Imperfect\ CSI}}&=M\log_2(1+\widehat{\eta}_{\mathrm{CI}})\\&=M\log_2\left(1+\frac{P\widehat{g}_{\mathrm{CI}}^2}{P\tau(1+\tau)+\sigma^2(1+\tau)^2}\right)\end{aligned} \tag{17}$$

3.3. Asymptotic performance analysis

In this subsection, we derive novel bounds regarding the asymptotic mean loss in sum rate and the achievable DoFs of linear downlink precoding. We do so with respect to the unified output SINRs of each user when CI is deployed at the BS in the presence of the imperfect CSI. By considering (10) and (17), the mean loss in sum rate can be shown as ΔR in (18) wherein ② is due to the fact that in (3), we defined $\tau=\beta\rho^{-\alpha}$. Consequently, the asymptotic mean loss in sum rate can be evaluated when the SNR goes to infinity and is therefore

$$\begin{aligned}\Delta R&=\mathbb{E}_{\mathbf{H}}\{R_{\mathtt{Perfect\ CSI}}\}-\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{R_{\mathtt{Imperfect\ CSI}}\}\\&=\mathbb{E}_{\mathbf{H}}\left\{M\log_2\left(1+\frac{Pg_{\mathrm{CI}}^2}{\sigma^2}\right)\right\}\\&\quad-\mathbb{E}_{\mathbf{H},\widehat{\mathbf{H}}}\left\{M\log_2\left(1+\frac{P\widehat{g}_{\mathrm{CI}}^2}{P\tau(1+\tau)+\sigma^2(1+\tau)^2}\right)\right\}\\&=\mathbb{E}_{\mathbf{H},\widehat{\mathbf{H}}}\left\{M\log_2\left(\frac{\left(Pg_{\mathrm{CI}}^2+\sigma^2\right)\left(P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)}{\sigma^2\left(P\widehat{g}_{\mathrm{CI}}^2+P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)}\right)\right\}\\&\overset{②}{=}\mathbb{E}_{\mathbf{H},\widehat{\mathbf{H}}}\\&\left\{M\log_2\left(\frac{\left(Pg_{\mathrm{CI}}^2+\sigma^2\right)\left(\beta P^{1-\alpha}\sigma^{2\alpha}(1+\beta P^{-\alpha}\sigma^{2\alpha})+\sigma^2(1+\beta P^{-\alpha}\sigma^{2\alpha})^2\right)}{\sigma^2\left(P\widehat{g}_{\mathrm{CI}}^2+\beta P^{1-\alpha}\sigma^{2\alpha}(1+\beta P^{-\alpha}\sigma^{2\alpha})+\sigma^2(1+\beta P^{-\alpha}\sigma^{2\alpha})^2\right)}\right)\right\}\end{aligned} \tag{18}$$

given by

$$\lim_{P\to\infty}\Delta R=\begin{cases}\infty & 0\le\alpha<1\\ C & \alpha=1\\ 0 & 1<\alpha\end{cases} \tag{19}$$

where $0<C<\infty$ is a constant which its value is given in the following theorem:

Theorem 1. *In reciprocal channels, when the error variance scales with the inverse of SNR, i.e., when $\alpha=1$, the asymptotic mean loss in sum rate is equal to $C=M\log_2(1+\beta)$.*

Proof. By considering (18), we have

$$\begin{aligned}\lim_{\substack{\alpha=1\\P\to\infty}}\Delta R&=\mathbb{E}_{\mathbf{H},\widehat{\mathbf{H}}}\left\{M\log_2\left(\frac{(1+\beta)g_{\mathrm{CI}}^2}{\widehat{g}_{\mathrm{CI}}^2}\right)\right\}\\&=\mathbb{E}_{\mathbf{H},\widehat{\mathbf{H}}}\left\{M\log_2\left(\frac{(1+\beta)\mathrm{Tr}\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]}{\mathrm{Tr}\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right]}\right)\right\}\end{aligned} \tag{20}$$

Due to the fact that $\lim_{\substack{\alpha=1\\P\to\infty}}\tau=\lim_{P\to\infty}\beta(P/\sigma^2)^{-1}=0$, and with respect to (2), we have

$$\lim_{\substack{\alpha=1\\P\to\infty}}\mathrm{Tr}\left[\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\right)^{-1}\right]=\mathrm{Tr}\left[\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}\right)^{-1}\right] \tag{21}$$

and consequently

$$C=\lim_{\substack{\alpha=1\\P\to\infty}}\Delta R=M\log_2(1+\beta)\quad\text{bits per channel use} \tag{22}$$

Eq. (22) implies that in reciprocal channels, although the full DoFs can be achieved, there is a nonzero constant gap between the curve representing the perfect CSI and the one denoting the imperfect CSI at high SNRs. This gap is equal to C. Since at high SNRs, each of the sum-rate curves has a slope of $M/3$ in units of bits per channel use per dB, the rate offset C, i.e., the vertical offset between the curve representing the perfect CSI and the one denoting the imperfect CSI case of $\alpha=1$, can be translated into a power offset, i.e., a horizontal offset, as follows:

$$\Delta\rho\Big|_{\substack{\alpha=1\\P\to\infty}}=\frac{3}{M}\Delta R\Big|_{\substack{\alpha=1\\P\to\infty}}=3\log_2(1+\beta)\ \mathrm{dB} \tag{23}$$

Eq. (23) implies that in reciprocal channels, we should increase the transmit power by $\Delta\rho$ dB to achieve the same sum rates as in the case of perfect CSI. As seen, unlike C in (22) which depends on both M and β, $\Delta\rho$ is only dependent on β.

Now that we have established bounds on asymptotic mean loss in sum rate, in Eq. (24), it is revealed that when $0\le\alpha<1$, an α fraction of the total DoF, i.e., $\alpha D_{\mathtt{Perfect\ CSI}}$ DoF, is achievable, where $D_{\mathtt{Perfect\ CSI}}$ is defined in (11).

Notice $0<\alpha<1$ reflects the scenario in which feedback power is much smaller than feedforward power. Therefore, the BS can reciprocally learn the forward link, but instead of full DoF, only an α fraction of that, i.e., αM, is achievable.

$$D_{\mathtt{Imperfect\ CSI}}=\lim_{P\to\infty}\frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\{R_{\mathtt{Imperfect\ CSI}}\}}{\log_2 P}$$

$$
\begin{aligned}
&= \lim_{P\to\infty} \frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(1+\frac{\widehat{g}_{\mathrm{CI}}^2 P}{P\tau(1+\tau)+\sigma^2(1+\tau)^2}\right)\right\}}{\log_2 P} \\
&= \lim_{P\to\infty} \frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(\widehat{g}_{\mathrm{CI}}^2 P+P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)\right\}}{\log_2 P} \\
&- \lim_{P\to\infty} \frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)\right\}}{\log_2 P} \\
&\geq \lim_{P\to\infty} \frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(\widehat{g}_{\mathrm{CI}}^2 P\right)\right\}}{\log_2 P} \\
&- \lim_{P\to\infty} \frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(P\tau(1+\tau)+\sigma^2(1+\tau)^2\right)\right\}}{\log_2 P} = M - \lim_{P\to\infty} \\
&\frac{\mathbb{E}_{\mathbf{H}|\widehat{\mathbf{H}}}\left\{M\log_2\left(\beta P^{1-\alpha}\sigma^{2\alpha}\left(1+\beta P^{-\alpha}\sigma^{2\alpha}\right)+\sigma^2\left(1+\beta P^{-\alpha}\sigma^{2\alpha}\right)^2\right)\right\}}{\log_2 P} \\
&= \begin{cases} M & 1\leq\alpha \\ \alpha M & 0\leq\alpha<1 \end{cases}
\end{aligned} \tag{24}
$$

Remark 1. Note that the results of (24) are inherently related to those in (19). For example, for the case of $\alpha=0$, i.e., the finite-rate feedback, while (24) implies that the achievable DoF is equal to zero, (19) indicates that for this case and by increasing the SNR, the asymptotic mean loss in sum rate is unboundedly increasing. Also (24) implies that when $\alpha\geq 1$, full DoF is achievable, that is, the asymptotic mean loss in sum rate is constant. In this case, (19) implies that when $\alpha=1$, this asymptotic mean loss converges to a non-zero constant whereas for $\alpha>1$, it tends to zero.

4. Standard RCI

In this section, we evaluate the performance of the standard RCI precoding [5] by deriving the output SINR of each user when it is deployed at the BS with the knowledge of the imperfect CSI. Note that the output SINR analysis of the standard RCI has been addressed in [5] where the derived formula just meant for the perfect CSI and is also dependent on the eigenvalues of $\mathbf{H}\mathbf{H}^{\mathrm{H}}$. However, in this paper, we derive the output SINR of each user based on a different approach which makes the RCI precoding especially amenable to the performance analysis subject to the imperfect CSI.

Under the assumption of the perfect CSI, the RCI precoder is defined as

$$
\mathbf{\Psi}_{\mathrm{RCI}} = \mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}+\varepsilon\mathbf{I}\right)^{-1} \tag{25}
$$

where $\varepsilon = M\rho^{-1}$ is the regularization parameter [5]. In this case, the scaling factor g can be shown as [27,28]

$$
g_{\mathrm{RCI}} = \frac{1}{\sqrt{\mathrm{Tr}\left[\mathbf{H}\mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{H}^{\mathrm{H}}+\varepsilon\mathbf{I}\right)^{-2}\right]}} \tag{26}
$$

By considering the fact that only the imperfect channel estimate $\widehat{\mathbf{H}}$ is available at the BS, the transmitted signal can be shown as

$$
\widehat{\mathbf{s}}_{\mathrm{RCI}} = \widehat{g}_{\mathrm{RCI}}\widehat{\mathbf{\Psi}}_{\mathrm{RCI}}\mathbf{c} \tag{27}
$$

such that the precoding matrix is defined as

$$
\widehat{\mathbf{\Psi}}_{\mathrm{RCI}} = \widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}+\varepsilon\mathbf{I}\right)^{-1} \tag{28}
$$

and the scaling factor is equal to

$$
\widehat{g}_{\mathrm{RCI}} = \frac{1}{\sqrt{\mathrm{Tr}\left[\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}+\varepsilon\mathbf{I}\right)^{-2}\right]}} \tag{29}
$$

Accordingly, the received signal of all users can be collectively represented by

$$
\begin{aligned}
\widehat{\mathbf{y}}_{\mathrm{RCI}} &= \sqrt{P}\widehat{g}_{\mathrm{RCI}}\mathbf{H}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}+\varepsilon\mathbf{I}\right)^{-1}\mathbf{c}+\mathbf{z} \\
&= \sqrt{P}\widehat{g}_{\mathrm{RCI}}\mathbf{H}\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}^{\mathrm{H}}\mathbf{c}+\mathbf{z}
\end{aligned} \tag{30}
$$

Let $\widehat{\mathbf{h}}_\ell \in \mathbb{C}^{1\times N}$ denote the ℓth row of $\widehat{\mathbf{H}}$ and $\widehat{\mathbf{H}}_\ell \in \mathbb{C}^{(M-1)\times N}$ designate the submatrix obtained by striking $\widehat{\mathbf{h}}_\ell$ out of $\widehat{\mathbf{H}}$. The received signal at the ℓth user is then given by Eq. (31) wherein ③ follows from (4), and we further considered $\mathbf{c}_\ell$ as a subvector obtained by removing c_ℓ from $\mathbf{c}$:

$$
\begin{aligned}
\widehat{y}_\ell &= \sqrt{P}\widehat{g}_{\mathrm{RCI}}\mathbf{h}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}^{\mathrm{H}}\mathbf{c}+z_\ell = \sqrt{P}\widehat{g}_{\mathrm{RCI}}\mathbf{h}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}c_\ell \\
&+ \sqrt{P}\widehat{g}_{\mathrm{RCI}}\mathbf{h}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\mathbf{c}_\ell+z_\ell \overset{③}{=} \underbrace{\sqrt{P}\frac{\widehat{g}_{\mathrm{RCI}}}{1+\tau}\widehat{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}c_\ell}_{\text{desired term}} \\
&\underbrace{+\sqrt{P}\frac{\widehat{g}_{\mathrm{RCI}}}{1+\tau}\widehat{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\mathbf{c}_\ell+\sqrt{P}\widehat{g}_{\mathrm{RCI}}\check{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}^{\mathrm{H}}\mathbf{c}+z_\ell}_{\text{interference plus noise term}}
\end{aligned} \tag{31}
$$

Note that based on the matrix inverse lemma,[1] we have

$$
\left(\widehat{\mathbf{H}}^{\mathrm{H}}\widehat{\mathbf{H}}+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}} = \frac{\left(\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}}{1+\widehat{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell+\varepsilon\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}} \tag{32}
$$

[1] If $\mathbf{x}$ is a row vector, then $\left(\mathbf{A}+\mathbf{x}^{\mathrm{H}}\mathbf{x}\right)^{-1}\mathbf{x}^{\mathrm{H}} = \frac{\mathbf{A}^{-1}\mathbf{x}^{\mathrm{H}}}{1+\mathbf{x}\mathbf{A}^{-1}\mathbf{x}^{\mathrm{H}}}$ [36].

In this case the desired signal power is equal to

$$E_{\text{desired signal}} = P\left[\frac{\widehat{g}_{\text{RCI}} A_\ell}{(1+\tau)(1+A_\ell)}\right]^2 \tag{33}$$

where $A_\ell = \widehat{\mathbf{h}}_\ell \left(\widehat{\mathbf{H}}_\ell^{\text{H}} \widehat{\mathbf{H}}_\ell + \varepsilon \mathbf{I}\right)^{-1} \widehat{\mathbf{h}}_\ell^{\text{H}}$. With respect to Lemma 1 and by taking the expectation over $\mathbf{c}$ and $\check{\mathbf{H}}$, the power of the interference term can be written as

$$E_{\text{interference}} = P\left[\frac{\widehat{g}_{\text{RCI}}}{(1+\tau)(1+A_\ell)}\right]^2 B_\ell + G \tag{34}$$

where

$$B_\ell = \widehat{\mathbf{h}}_\ell \left(\widehat{\mathbf{H}}_\ell^{\text{H}} \widehat{\mathbf{H}}_\ell + \varepsilon \mathbf{I}\right)^{-1} \widehat{\mathbf{H}}_\ell^{\text{H}} \widehat{\mathbf{H}}_\ell \left(\widehat{\mathbf{H}}_\ell^{\text{H}} \widehat{\mathbf{H}}_\ell + \varepsilon \mathbf{I}\right)^{-1} \widehat{\mathbf{h}}_\ell^{\text{H}}$$

and by considering (29), we have

$$G = P\widehat{g}_{\text{RCI}}^2 \, \mathbb{E}_{\check{\mathbf{h}}_\ell, \mathbf{c}} \left\{ \check{\mathbf{h}}_\ell \left(\widehat{\mathbf{H}}^{\text{H}} \widehat{\mathbf{H}} + \varepsilon \mathbf{I}\right)^{-1} \widehat{\mathbf{H}}^{\text{H}} \mathbf{c} \, \mathbf{c}^{\text{H}} \widehat{\mathbf{H}} \left(\widehat{\mathbf{H}}^{\text{H}} \widehat{\mathbf{H}} + \varepsilon \mathbf{I}\right)^{-1} \check{\mathbf{h}}_\ell^{\text{H}} \right\} = \frac{P\tau}{1+\tau} \tag{35}$$

Consequently, the output SINR of the ℓth user can be shown as

$$\widehat{\eta}_{\text{RCI}} = \frac{E_{\text{desired signal}}}{E_{\text{interference}} + \sigma^2} = \frac{\widehat{g}_{\text{RCI}}^2 A_\ell^2 P}{\widehat{g}_{\text{RCI}}^2 B_\ell P + P\tau(1+\tau)(1+A_\ell)^2 + \sigma^2(1+\tau)^2(1+A_\ell)^2} \tag{36}$$

5. Adaptive RCI

Subject to the perfect CSI, the standard RCI precoding outperforms the CI precoding; however, in the presence of imperfect CSI, its comparative improvement to the CI deteriorates. In other words, the standard RCI is more sensitive to the CSI mismatch than the CI precoding is. Therefore, in this section and by deriving an appropriate regularization parameter, we propose an adaptive RCI that outperforms the standard RCI under imperfect CSI. To do so, we further assume that the BS knows the variance of the channel estimation error, i.e., τ, which is possible to be known in advance, as discussed in Section 2.2.

We obtain the adaptive RCI precoder by using the following minimum-mean-square-error (MMSE) criterion:

$$\min_{\widehat{\boldsymbol{\Psi}}} \mathbb{E}\left\{ \left\| \sqrt{P}\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c} + f\mathbf{z} - \sqrt{P}\mathbf{c} \right\|_2^2 \right\} \tag{37}$$

where

$$f = \frac{1+\tau}{\tilde{g}} = (1+\tau)\sqrt{\text{Tr}\left[\widehat{\boldsymbol{\Psi}}^{\text{H}} \widehat{\boldsymbol{\Psi}}\right]} \tag{38}$$

and $\tilde{g}$ is the scaling factor. The inclusion of f in (37) is due to the fact that in all precoding schemes, each MT uses the knowledge of the scaling factor to estimate the transmitted data, which makes the system be equivalent to the one that has a scaled noise variance, and consequently, this effect can be reflected through a multiplicative factor like f. In other words, at transmit side, the transmitted signals are scaled by $\tilde{g}$ to meet the power constraints; consequently, at receive side and by considering (14) or (31), to have an unbiased detection, the received signals should be scaled back by $(1+\tau)/\tilde{g}$, which further appears as a multiplicative factor for the noise.

The objective function in (37) can then be shown as

$$\begin{aligned} F &= \mathbb{E}\left\{ \text{Tr}\left[\left(\sqrt{P}\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c} + f\mathbf{z} - \sqrt{P}\mathbf{c} \right) \cdot \left(\sqrt{P}\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c} + f\mathbf{z} - \sqrt{P}\mathbf{c} \right)^{\text{H}} \right] \right\} \\ &= \mathbb{E}\left\{ \text{Tr}\left[P\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c}\mathbf{c}^{\text{H}}\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}} + f^2\mathbf{z}\mathbf{z}^{\text{H}} - P\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c}\mathbf{c}^{\text{H}} - P\mathbf{c}\mathbf{c}^{\text{H}}\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}} \right.\right. \\ &\quad \left.\left. + P\mathbf{c}\mathbf{c}^{\text{H}} + \sqrt{P}f\mathbf{H}\widehat{\boldsymbol{\Psi}}\mathbf{c}\mathbf{z}^{\text{H}} + \sqrt{P}f\mathbf{z}\mathbf{c}^{\text{H}}\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}} \right.\right. \\ &\quad \left.\left. - \sqrt{P}f\mathbf{c}\mathbf{z}^{\text{H}} - \sqrt{P}f\mathbf{z}\mathbf{c}^{\text{H}} \right] \right\} \\ &\overset{④}{=} P\text{Tr}\left[\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{H}\widehat{\boldsymbol{\Psi}}\right] + M\sigma^2(1+\tau)^2\text{Tr}\left[\widehat{\boldsymbol{\Psi}}^{\text{H}}\widehat{\boldsymbol{\Psi}}\right] - P\text{Tr}\left[\mathbf{H}\widehat{\boldsymbol{\Psi}}\right] \\ &\quad - P\text{Tr}\left[\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}}\right] + PM \end{aligned} \tag{39}$$

where ④ follows the fact that the data and noise are considered to be independent of $\widehat{\mathbf{H}}$ and $\check{\mathbf{H}}$. To obtain the adaptive precoder, we can differentiate F with respect to $\widehat{\boldsymbol{\Psi}}$ by first considering the following assumptions [29,31]:

1. $\widehat{\boldsymbol{\Psi}}$ and $\widehat{\boldsymbol{\Psi}}^{\text{H}}$ are treated as independent variables.

2. $\dfrac{\partial \text{Tr}\left[\mathbf{A}\widehat{\boldsymbol{\Psi}}\right]}{\partial \widehat{\boldsymbol{\Psi}}} = \dfrac{\partial \text{Tr}\left[\widehat{\boldsymbol{\Psi}}\mathbf{A}\right]}{\partial \widehat{\boldsymbol{\Psi}}} = \mathbf{A}.$

Following the preceding assumptions, the differentiation of F with respect to $\widehat{\boldsymbol{\Psi}}$ gives

$$\begin{aligned} \frac{\partial F}{\partial \widehat{\boldsymbol{\Psi}}} &= P\widehat{\boldsymbol{\Psi}}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{H} + M\sigma^2(1+\tau)^2\widehat{\boldsymbol{\Psi}}^{\text{H}} - P\mathbf{H} \\ &\overset{⑤}{=} P\widehat{\boldsymbol{\Psi}}^{\text{H}} \left(\frac{\widehat{\mathbf{H}}}{1+\tau} + \check{\mathbf{H}} \right)^{\text{H}} \left(\frac{\widehat{\mathbf{H}}}{1+\tau} + \check{\mathbf{H}} \right) \\ &\quad + M\sigma^2(1+\tau)^2\widehat{\boldsymbol{\Psi}}^{\text{H}} - P\left(\frac{\widehat{\mathbf{H}}}{1+\tau} + \check{\mathbf{H}} \right) \\ &= P\boldsymbol{\Psi}^{\text{H}} \left[\frac{\widehat{\mathbf{H}}^{\text{H}}\widehat{\mathbf{H}}}{(1+\tau)^2} + \check{\mathbf{H}}^{\text{H}}\check{\mathbf{H}} + \frac{\widehat{\mathbf{H}}^{\text{H}}\check{\mathbf{H}} + \check{\mathbf{H}}^{\text{H}}\widehat{\mathbf{H}}}{1+\tau} \right] \\ &\quad + M\sigma^2(1+\tau)^2\widehat{\boldsymbol{\Psi}}^{\text{H}} - \frac{P}{1+\tau}\widehat{\mathbf{H}} - P\check{\mathbf{H}} \end{aligned} \tag{40}$$

where ⑤ follows from (4). The adaptive precoder can then be found by setting $\partial F/\partial \widehat{\boldsymbol{\Psi}}$ equal to zero and taking the expectation over the auxiliary random matrix $\check{\mathbf{H}}$. First, note that we have

$$\mathbb{E}\left\{\check{\mathbf{H}}\right\} = 0 \tag{41a}$$

$$\mathbb{E}\left\{\check{\mathbf{H}}^{\text{H}}\check{\mathbf{H}}\right\} \overset{⑥}{=} \frac{M\tau}{1+\tau}\mathbf{I} \tag{41b}$$

where ⑥ follows Lemma 1. Consequently, we can represent the precoding matrix as

$$\mathbb{E}_{\check{\mathbf{H}}}\left\{\frac{\partial F}{\partial \widehat{\boldsymbol{\Psi}}}\right\} = \mathbf{0} \Longrightarrow \widehat{\boldsymbol{\Psi}} = \widehat{\mathbf{H}}^{\text{H}} \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\text{H}} + \widehat{\varepsilon}\,\mathbf{I} \right)^{-1} \tag{42}$$

where the regularization parameter $\widehat{\varepsilon}$ can now be expressed as

$$\widehat{\varepsilon} = M(1+\tau)\left(\tau + \rho^{-1}(1+\tau)^3\right) \tag{43}$$

Remark 2. Since the Hessian matrix of the MSE objective function is positive definite, the expression in (42) is a global minimizer for the considered MMSE optimization problem. This implies the optimality of the derived regularization parameter in (43). Note that by setting $\tau=0$, $\widehat{\varepsilon}$ boils down to $\varepsilon=M\rho^{-1}$ which is the appropriate regularization parameter under perfect CSI that achieves the best tradeoff between noise and multiuser interference [5].

Therefore, for the proposed adaptive RCI, the transmitted signal from the BS can be shown as

$$\widehat{\mathbf{s}}_{\texttt{adaptive RCI}}=\tilde{g}\widehat{\mathbf{\Psi}}\mathbf{c} \tag{44}$$

where $\widehat{\mathbf{\Psi}}$ is defined in (42), and $\tilde{g}$ is the scaling factor which can be represented as

$$\tilde{g}=\frac{1}{\sqrt{\mathrm{Tr}\left[\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\mathrm{H}}+\widehat{\varepsilon}\,\mathbf{I}\right)^{-2}\right]}} \tag{45}$$

Similar to the standard RCI, it is straightforward to show that the output SINR of the ℓth user based on the adaptive RCI can now be expressed as

$$\tilde{\eta}_{\texttt{RCI}}=\frac{\tilde{g}^2\widehat{A}_\ell^2P}{\tilde{g}^2\widehat{B}_\ell P+P\tau(1+\tau)\left(1+\widehat{A}_\ell\right)^2+\sigma^2(1+\tau)^2\left(1+\widehat{A}_\ell\right)^2} \tag{46}$$

where $\widehat{A}_\ell=\widehat{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell+\widehat{\varepsilon}\,\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}$ and

$\widehat{B}_\ell=\widehat{\mathbf{h}}_\ell\left(\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell+\widehat{\varepsilon}\,\mathbf{I}\right)^{-1}\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell\left(\widehat{\mathbf{H}}_\ell^{\mathrm{H}}\widehat{\mathbf{H}}_\ell+\widehat{\varepsilon}\,\mathbf{I}\right)^{-1}\widehat{\mathbf{h}}_\ell^{\mathrm{H}}$

such that $\widehat{\varepsilon}$ is defined in (43).

Remark 3. The output SINR of each user due to the standard RCI with perfect CSI can be easily obtained by setting $\tau=0$ and replacing $\widehat{\mathbf{H}}$ with $\mathbf{H}$ in $\widehat{A}_\ell$, $\widehat{B}_\ell$ and $\tilde{g}$ in Eq. (46). In other words, the adaptive RCI is a generalized and optimized version of the standard RCI in [5] without introducing any extra computational complexity.

Remark 4. Note that although the derived bounds in (19)–(24) are based on the output SINR of CI precoding, they are likewise applicable to the case of RCI. This is due to the fact that the output SINR of MMSE-based equalizers, conditioned on the channel realization, is asymptotically equal to that of ZF-based equalizers plus a gap [37], i.e., $\eta_{\texttt{RCI}}=\eta_{\texttt{CI}}+\varphi$ where φ is the aforementioned gap. This implies that at high SNRs, φ becomes negligible compared to both $\eta_{\texttt{RCI}}$ and $\eta_{\texttt{CI}}$.

6. Numerical results

In this section, by using simulation results, we substantiate the analytically derived bounds in (19)–(24). We also demonstrate the superior performance achieved by adaptive RCI compared to standard RCI.

We evaluate the sum rates as [5,38]

$$\sum_{\ell=1}^{M}\log_2(1+\texttt{SINR}_\ell) \tag{47}$$

where $\texttt{SINR}_\ell$ denotes the output SINR of the ℓth user. For instance, in the case of adaptive RCI, $\texttt{SINR}_\ell=\tilde{\eta}_{\texttt{RCI}}$ where $\tilde{\eta}_{\texttt{RCI}}$ is defined in (46).

In the interest of verifying the accuracy of the derived output SINR in (46) and to analytically evaluate the BER of adaptive RCI precoding, we utilize the following formula within Fig. 2, which is a reliable criterion to analytically evaluate the BER of each user when $\mathcal{M}$-QAM constellation with Gray code bit mapping is used [39]

$$\texttt{BER}_{\mathcal{M}-\texttt{QAM}}\cong\frac{4}{\log_2\mathcal{M}}\left(1-\frac{1}{\sqrt{\mathcal{M}}}\right)\times\sum_{i=1}^{2}Q\left((2i-1)\sqrt{\frac{3\times\texttt{SINR}_\ell}{\mathcal{M}-1}}\right) \tag{48}$$

wherein

$$Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}\mathrm{d}t \tag{49}$$

In Fig. 2, we assume communications under 64-QAM signaling with Gray code bit mapping and $M=N=8$. Simulated results are due to counting the number of occurred errors in received bits when the transmitted signals are based on what is expressed in (44). As revealed, both analytical and simulated results are in excellent agreement, which verify the validity of the derived output SINR in (46). This can be similarly used to verify the validity of the derived SINRs of the CI and standard RCI as well, though the corresponding curves are omitted for the sake of compactness.

In Fig. 3, we certify the aforementioned bounds using CI precoding subject to different CSI qualities and for the case $M=5,N=8$. With respect to the fact that the achievable DoF under perfect CSI is equal to 5, the following performance trends are observed:

- $\alpha>1$: While (19) indicates that the asymptotic mean loss in sum rate is equal to zero, (24) denotes that the full DoF, i.e., 5 DoFs, should be achievable. All these bounds are certified where the corresponding curve overlaps with the one representing the perfect CSI at high SNRs.
- $\alpha=1$: While (19) indicates that the asymptotic mean loss in sum rate is equal to a nonzero finite constant, (24) denotes that the full DoF, i.e., 5 DoFs, should be achievable. These bounds are also certified where the corresponding curve has the same slope as in the case of perfect CSI, and consequently there is a nonzero constant gap between them.
- For $\beta=10,\alpha=1$ and based on (22), we expect that the asymptotic mean loss in sum rate should be 17.3 bits per channel use, which is verified by the depicted results in Fig. 3. Plus, based on (23), we expect that in the case of $\beta=10,\alpha=1$ and to achieve the same sum rate as in the case of perfect CSI, we should increase the transmit power by 10.4 dB, which is again certified in Fig. 3.
- $0\le\alpha<1$: In this case, Eq. (19) indicates that the asymptotic mean loss in sum rate is unboundedly increasing with SNR. This is again certified in Fig. 3, such that when SNR gets larger, the gap between the corresponding curves and that of the perfect CSI becomes wider. Also based on (24), we expect that an

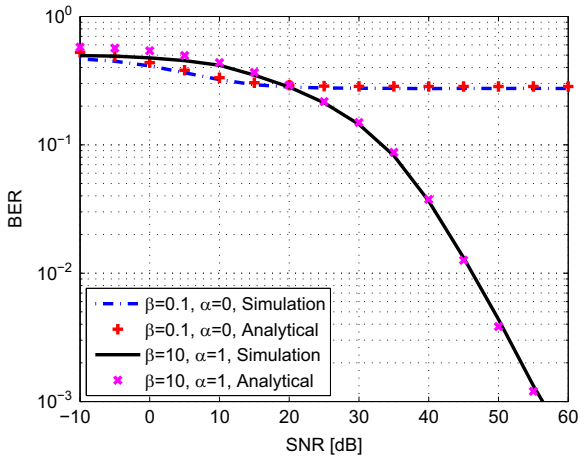

Fig. 2. Comparison between analytical and simulated BERs of adaptive RCI precoding under 64-QAM signaling for $M=N=8$ when $\beta=0.1,\ \alpha=0$, and $\beta=10,\ \alpha=1$.

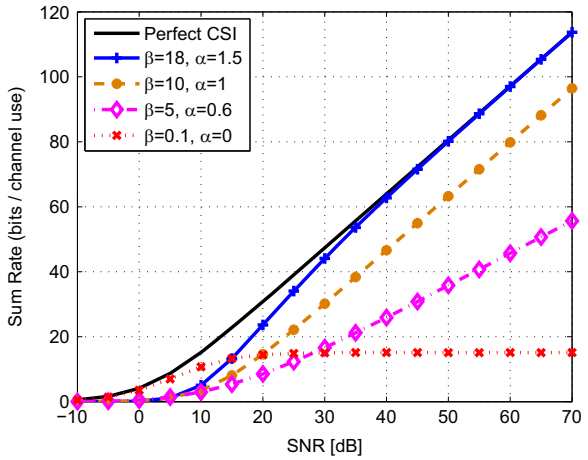

Fig. 3. Sum rates of CI precoding for $M=5,\ N=8$ and under different CSI qualities.

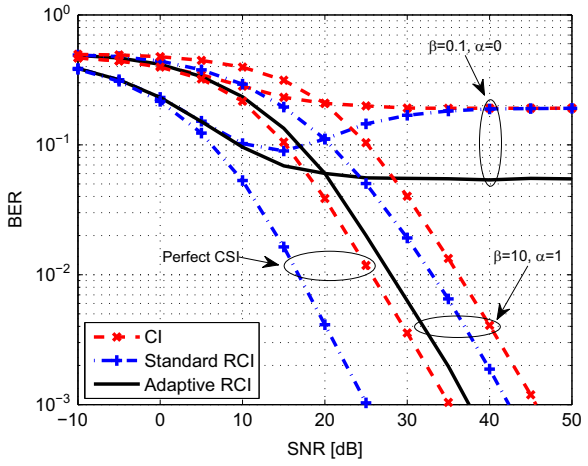

Fig. 4. BER under QPSK signaling for $M=N=8$ and for the cases $\beta=0.1,\ \alpha=0$, and $\beta=10,\ \alpha=1$.

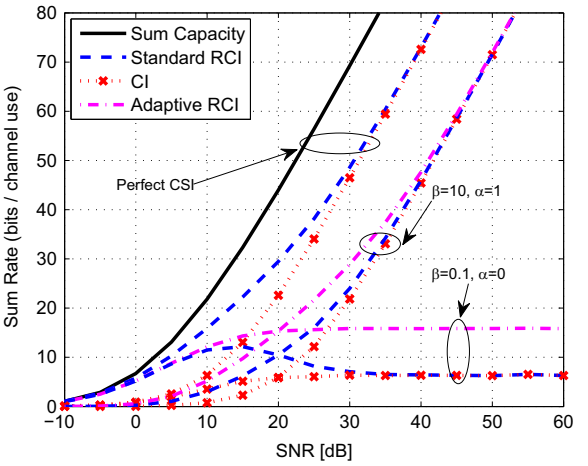

Fig. 5. Sum rates for $M=N=8$, when $\beta=0.1,\ \alpha=0$, and $\beta=10,\ \alpha=1$.

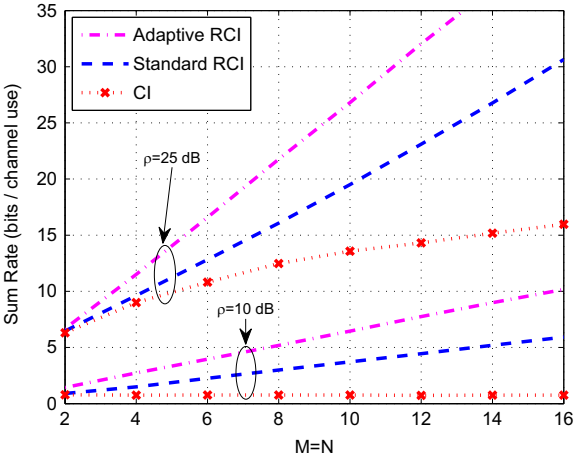

Fig. 6. Sum rates of adaptive and standard RCI compared to those of CI as a function of M and N, at SNRs of 10 dB and 25 dB, and for the case of $\beta=10,\ \alpha=1$.

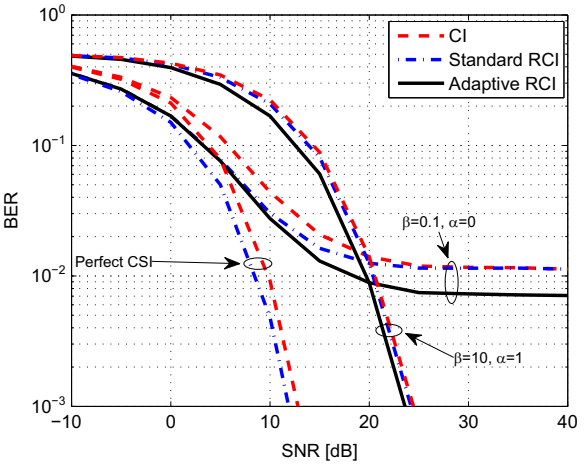

Fig. 7. BER under QPSK signaling for $M=5,\ N=8$ when $\beta=0.1,\ \alpha=0$, and $\beta=10,\ \alpha=1$.

α fraction of the total DoF should be achievable. By considering the slopes of the curves in the same figure, while for $\alpha = 0.6$ the achievable DoF is 3, for the case $\alpha = 0$, it is equal to zero.

Although the promised improvement of adaptive RCI over standard RCI and CI can be gleaned for various values of α, in Figs. 4–7 and without loss of generality, we just focus on two representative cases: $\alpha = 0$ (which mimics the CSI feedback scenario), and $\alpha = 1$ (which imitates the reciprocal channels). More specifically and with respect to the error variance τ defined in (3), we consider two cases: $\beta = 10, \alpha = 1$ and $\beta = 0.1, \alpha = 0$. We also assume that $M = N = 8$ unless stated otherwise.

Fig. 4 illustrates the BER of CI and RCI under QPSK signaling. As demonstrated, the proposed adaptive RCI achieves better BER than standard RCI. For example, for the case of $\alpha = 1$, adaptive RCI attains nearly 6 dB and 9 dB gain compared to standard RCI and CI, respectively, to achieve the BER of 10^{-3}.

In Fig. 5, the sum rates of CI and RCI under perfect and imperfect CSI are depicted. As shown, for different values of α, the adaptive RCI achieves higher sum rates than the standard RCI does, e.g., for the case of CSI feedback ($\alpha = 0$) and at high SNRs, while standard RCI achieves the same sum rate as CI does, adaptive RCI yields nearly 10 bits per channel use gain in sum rate. Note that when $\alpha = 1$, we expect that the achievable DoF should be the same as the one in the case of perfect CSI. This is confirmed in Fig. 5 where it can be seen that the curves related to $\alpha = 1$ have the same slope as the ones pertained to the perfect CSI, at high SNRs.

By considering Figs. 4 and 5, one interesting observation is that for the CSI feedback ($\alpha = 0$), while the performance trend of the standard RCI is nonmonotonic, that of the adaptive RCI is monotonic. This nonmonotonic behavior of the standard RCI precoding is related to the case of digital feedback where at high SNRs, the system becomes interference-limited such that by increasing the operational SNR, the performance trend first becomes improved at low-to-intermediate SNRs, but suddenly deteriorates at a saddle point, and eventually becomes saturated at high SNRs, which leads to a nonmonotonic behavior. For the adaptive RCI precoding, on the other hand, the performance trend does not suddenly deteriorate. This is due to its regularization parameter which is a function of the channel estimation error variance τ. Thus, in this case, the performance trend smoothly becomes saturated, which results in a monotonic behavior.

Fig. 6 depicts the average sum rates of linear precoders as a function of M and N, at $\rho = 10$ dB and $\rho = 25$ dB and in the case of reciprocal channels, i.e., $\alpha = 1$. As revealed, with increasing M, while the sum rate of CI does not linearly increase, those of standard and adaptive RCI do. Also adaptive RCI outperforms standard RCI and CI at both low and high SNRs such that the larger the M and N are, the more gain in sum rate can be gleaned by deploying adaptive RCI.

As mentioned earlier, the concept of regularization is most beneficial with equal number of transmit and receive antennas [5,24,27,28]. Nevertheless, in Fig. 7, we compare the BERs of the adaptive RCI with those of the standard RCI and CI when the number of antennas at the BS is more than the number of receive antennas. The results are depicted under QPSK signaling, $\beta = 10, \alpha = 1$ and $\beta = 0.1, \alpha = 0$, when $M = 5$ and $N = 8$. As revealed, even in this case, the adaptive RCI achieves better performance than standard RCI and CI do. For instance, when $\alpha = 1$, while the BER of standard RCI is the same as that of CI, adaptive RCI achieves 1 dB gain to reach the BER of 10^{-3}. However, for the case of $\alpha = 0$, while at high SNRs, standard RCI and CI achieve the same BERs, adaptive RCI achieves a distinguished performance such that there is a gap between the BER of adaptive RCI and those of the standard RCI and CI.

Finally, it is worthwhile to point out why we reckon that the adaptive RCI compensates the degraded performance of the standard RCI compared to CI under CSI mismatch. This can be clearly observed in Figs. 4, 5 and 7. For example, as revealed in Fig. 4, while under perfect CSI, standard RCI achieves nearly 10 dB gain compared to CI to reach the BER of 10^{-3}, under imperfect CSI, say $\alpha = 1$, this gain is nearly 4 dB. However, adaptive RCI tries to compensate this 6 dB loss in performance such that the achieved gain is now nearly 10 dB.

7. Conclusions

In this paper, we quantified the performance of linear precoders, namely CI and RCI, in the multiuser multiantenna downlink under a generalized CSI mismatch model where the variance of the CSI measurement error depends on the operational SNR. We derived novel bounds regarding the asymptotic mean loss in sum rate and the achievable DoF. For example, we showed that when this error variance scales with the inverse of the operational SNR, full DoF is achievable, and the asymptotic mean loss in sum rate is thus a nonzero finite constant which its exact value was derived. Then, this mean loss in sum rate was shown to be equivalent to a power loss. It was also demonstrated that under imperfect CSI, the comparative improvement of standard RCI to CI becomes negligible, which implies on the more sensitivity of standard RCI to imperfect CSI compared to CI precoding. Accordingly, we proposed an adaptive RCI by deriving an appropriate regularization parameter as a function of the error variance and without imposing any restrictions on the number of users and antennas. Simulation results showed that the adaptive RCI outperforms the standard RCI under CSI mismatch such that it compensates the degraded performance of standard RCI compared to CI without introducing any further computational complexity.

References

[1] H. Weingarten, Y. Steinberg, S. Shamai (Shitz), The capacity region of the Gaussian multiple-input multiple-output broadcast channel, IEEE Trans. Inf. Theory 52 (9) (2006) 3936–3964.
[2] B.M. Hochwald, C.B. Peel, A.L. Swindlehurst, A vector-perturbation technique for near-capacity multiantenna multiuser communication-part II: perturbation, IEEE Trans. Commun. 53 (3) (2005) 537–544.
[3] J. Maurer, J. Jaldén, D. Seethaler, G. Matz, Vector perturbation precoding revisited, IEEE Trans. Signal Process. 59 (1) (2011) 315–328.
[4] L. Sanguinetti, M. Morelli, Non-linear pre-coding for multiple-antenna multi-user downlink transmissions with different QoS requirements, IEEE Trans. Wirel. Commun. 6 (3) (2007) 852–856.

[5] C.B. Peel, B.M. Hochwald, A.L. Swindlehurst, A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization, IEEE Trans. Commun. 53 (1) (2005) 192–202.
[6] T. Yoo, A. Goldsmith, On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming, IEEE J. Sel. Areas Commun. 24 (3) (2006) 528–541.
[7] N. Jindal, MIMO broadcast channels with finite rate feedback, IEEE Trans. Inf. Theory 52 (11) (2006) 5045–5060.
[8] A.D. Dabbagh, D.J. Love, Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch, IEEE Trans. Commun. 56 (11) (2008) 1859–1868.
[9] H. Sung, S.-R. Lee, I. Lee, Generalized channel inversion methods for multiuser MIMO systems, IEEE Trans. Commun. 57 (11) (2009) 3489–3499.
[10] B. Song, M. Haardt, Effects of imperfect channel state information on achievable rates of precoded multi-user MIMO broadcast channels with limited feedback, in: Proceedings of IEEE International Conference on Communications (ICC), 2009, pp. 1–5.
[11] D.J. Love, R.W. Heath, W. Santipach, M.L. Honig, What is the value of limited feedback for MIMO channels? IEEE Commun. Mag. 42 (10) (2004) 54–59.
[12] D.J. Love, et al., An overview of limited feedback in wireless communication systems, IEEE J. Sel. Areas Commun. 26 (8) (2008) 1341–1365.
[13] T. Yoo, N. Jindal, A. Goldsmith, Multi-antenna downlink channels with limited feedback and user selection, IEEE J. Sel. Areas Commun. 25 (7) (2007) 1478–1491.
[14] P. Xia, G.B. Giannakis, Design and analysis of transmit-beamforming based on limited-rate feedback, IEEE Trans. Signal Process. 54 (5) (2006) 1853–1863.
[15] P. Ding, D.J. Love, M.D. Zoltowski, Multiple antenna broadcast channels with shape feedback and limited feedback, IEEE Trans. Signal Process. 55 (7) (2007) 3417–3428.
[16] N. Ravindran, N. Jindal, Multi-user diversity vs. accurate channel state information in MIMO downlink channels, IEEE Trans. Wirel. Commun. 11 (9) (2012) 3037–3046.
[17] T.L. Marzetta, B.M. Hochwald, Fast transfer of channel state information in wireless systems, IEEE Trans. Signal Process. 54 (4) (2006) 1268–1278.
[18] U. Salim, D. Gesbert, D. Slock, Combining training and quantized feedback in multiantenna reciprocal channels, IEEE Trans. Signal Process. 60 (3) (2012) 1383–1396.
[19] J. Shi, M. Ho, MIMO broadcast channels with channel estimation, in: Proceedings of IEEE International Conference on Communications (ICC), 2007, pp. 1042–1047.
[20] B. Zhou, et al., Impact of imperfect channel state information on TDD downlink multiuser MIMO system, in: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), 2011, pp. 1823–1828.
[21] R. Hunger, F.A. Dietrich, M. Joham, W. Utschick, Robust transmit zero-forcing filters, in: Proceedings of ITG Workshop on Smart Antennas, 2004, pp. 130–137.
[22] M. Kobayashi, N. Jindal, G. Caire, Training and feedback optimization for multiuser MIMO downlink, IEEE Trans. Commun. 59 (8) (2011) 2228–2240.
[23] G. Caire, N. Jindal, M. Kobayashi, N. Ravindran, M.I.M.O. Multiuser, Multiuser MIMO achievable rates with downlink training and channel state feedback, IEEE Trans. Inf. Theory 56 (6) (2010) 2845–2866.
[24] J. Zhang, C.-K. Wen, S. Jin, X. Gao, K.-K. Wong, Large system analysis of cooperative multi-cell downlink transmission via regularized channel inversion with imperfect CSIT, IEEE Trans. Wirel. Commun. 12 (10) (2013) 4801–4813.
[25] S. Wagner, R. Couillet, M. Debbah, D.T.M. Slock, Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback, IEEE Trans. Inf. Theory 58 (7) (2012) 4509–4537.
[26] V.K. Nguyen, J.S. Evans, Multiuser transmit beamforming via regularized channel inversion: a large system analysis, in: Proceedings of IEEE Global Communications Conference (GLOBECOM), 2008, pp. 1–4.
[27] R. Muharar, R. Zakhour, J. Evans, Optimal power allocation and user loading for multiuser MISO channels with regularized channel inversion, IEEE Trans. Commun. 61 (12) (2013) 5030–5041.
[28] S.M. Razavi, T. Ratnarajah, C. Masouros, Transmit-power efficient linear precoding utilizing known interference for the multiantenna downlink, IEEE Trans. Veh. Technol. 63 (9) (2014) 4383–4394.
[29] S.M. Razavi, T. Ratnarajah, Adaptively regularized phase alignment precoding for multiuser multiantenna downlink, IEEE Trans. Veh. Technol. 64 (10) (2015) 4863–4869.
[30] S.M. Razavi, T. Ratnarajah, Performance analysis of interference alignment under CSI mismatch, IEEE Trans. Veh. Technol. 63 (9) (2014) 4740–4748.
[31] S.M. Razavi, T. Ratnarajah, Adaptive LS and MMSE based beamformer design for multiuser MIMO interference channels, IEEE Trans. Veh. Technol. [Online]. Available: ⟨http://ieeexplore.ieee.org/arnumber=7008543⟩.
[32] T. Yoo, A. Goldsmith, Capacity and power allocation for fading MIMO channels with channel estimation error, IEEE Trans. Inf. Theory 52 (5) (2006) 2203–2214.
[33] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall, New Jersey, 1993.
[34] A.M. Tulino, S. Verdú, Random matrix theory and wireless communications, Found. Trends Commun. Inf. Theory 1 (1) (2004) 1–182.
[35] C. Wang, E.K.S. Au, R.D. Murch, W.H. Mow, R.S. Cheng, V. Lau, On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error, IEEE Trans. Wirel. Commun. 6 (3) (2007) 805–810.
[36] G.H. Golub, C.F. Van Loan, Matrix Computations, 3rd ed. The Johns Hopkins University Press, Baltimore, London, 1996.
[37] Y. Jiang, M.K. Varanasi, J. Li, Performance analysis of ZF and MMSE equalizers for MIMO systems: an in-depth study of the high SNR regime, IEEE Trans. Inf. Theory 57 (4) (2011) 2008–2026.
[38] M.R. McKay, I.B. Collings, A.M. Tulino, Achievable sum rate of MIMO MMSE receivers: a general analytic framework, IEEE Trans. Inf. Theory 56 (1) (2010) 396–410.
[39] J. Lu, K.B. Letaief, J.C.-I. Chuang, M.L. Liou, M-PSK and M-QAM BER computation using signal-space concepts, IEEE Trans. Commun. 47 (2) (1999) 181–184.